Adjoint Method Explained

Sarka Tukova

Structure of the talk

- Motivation
- Example
- Forward and Adjoint
- Limitations
- Store or Recalculate
- Continuous or Discrete?

Motivation - Sensitivity Analysis

Main objective:

- to quantify the effects of parameter variations on calculated results

Results in assessment of:

- influence
- importance
- dominance

Methods used:

- forward sensitivity - tangent linear model
- reverse sensitivity - adjoint model
- other

Motivation - Other

The adjoint model can also be used for:

- Error estimation
- Data assimilation - finding model solutions to best fit given data
- Calibration of model to data
- Optimisation

General Method

Define the model equation(s)

General Method

Define the model equation(s)

$$\downarrow$$

Define the cost function(s)

General Method

Define the model equation(s)

$$\downarrow$$

Define the cost function(s)

$$\downarrow$$

Derive the adjoint equation(s)

Example - Model Equation

Navier-Stokes Equations:

$$\frac{\partial \boldsymbol{u}}{\partial t} + (\boldsymbol{u} \cdot \nabla)\boldsymbol{u} + f\boldsymbol{k} \times \boldsymbol{u} = -\nabla p - \rho g \boldsymbol{k} + \nabla \cdot \tau$$

Example - Model Equation

Navier-Stokes Equations:

$$\frac{\partial \boldsymbol{u}}{\partial t} + (\boldsymbol{u} \cdot \nabla)\boldsymbol{u} + f\boldsymbol{k} \times \boldsymbol{u} = -\nabla p - \rho g\boldsymbol{k} + \nabla \cdot \tau$$

First two terms from 1D Navier-Stokes Equation:

$$F(u, \frac{\partial u}{\partial t}, \frac{\partial u}{\partial x}, m) = F(u, u_t, u_x, m) = \frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0$$

Example - Cost Function

Choose a cost function according to what you would like to test.

Here we will consider a cost function in a general form:

$$J = \int_t \int_\Omega j(x, u, m) dx dt$$

where m are the control input variables.

Example - Euler-Lagrange Method

To derive the adjoint equation we use the Euler-Lagrange method:

Example - Euler-Lagrange Method

To derive the adjoint equation we use the Euler-Lagrange method:

- Construct the Lagrange functional from the weak form of Navier-Stokes equations and the cost function:

$$L(x,u,m) = J(x,u,m) + \int_t \int_\Omega F u^* dx dt$$

where u^* is the adjoint variable - independent of m!

Example - Differentiating

The sensitivity of the model to the control input variables, m, is found by differentiating L with respect to m:

Example - Differentiating

The sensitivity of the model to the control input variables, m, is found by differentiating L with respect to m:

$$\frac{dL(x,u,m)}{dm} = \frac{dJ(x,u,m)}{dm} + \int_t \int_\Omega \frac{dF(u,u_t,u_x,m)}{dm} u^* dx dt$$

Example - Differentiating

The sensitivity of the model to the control input variables, m, is found by differentiating L with respect to m:

$$\frac{dL(x,u,m)}{dm} = \frac{dJ(x,u,m)}{dm} + \int_t \int_\Omega \frac{dF(u,u_t,u_x,m)}{dm} u^* dx dt$$

$$\frac{dJ}{dm} = \int_t \int_\Omega \frac{dj}{dm} dx dt = \int_t \int_\Omega j_m + j_u u_m dx dt$$

Example - Differentiating

The sensitivity of the model to the control input variables, m, is found by differentiating L with respect to m:

$$\frac{dL(x,u,m)}{dm} = \frac{dJ(x,u,m)}{dm} + \int_t \int_\Omega \frac{dF(u,u_t,u_x,m)}{dm} u^* dxdt$$

$$\frac{dJ}{dm} = \int_t \int_\Omega \frac{dj}{dm} dxdt = \int_t \int_\Omega j_m + j_u u_m dxdt$$

$$\int_t \int_\Omega \frac{dF}{dm} u^* dxdt = \int_t \int_\Omega (F_m + F_u u_m + F_{u_t} u_{tm} + F_{u_x} u_{xm}) u^* dxdt$$

Example - Integration by Parts

$$\int_t \int_\Omega \frac{dF}{dm} u^* dx dt = \int_t \int_\Omega (F_m + F_u u_m + F_{u_t} u_{tm} + F_{u_x} u_{xm}) u^* dx dt$$

Example - Integration by Parts

$$\int_t \int_\Omega \frac{dF}{dm} u^* dxdt = \int_t \int_\Omega (F_m + F_u u_m + F_{u_t} u_{tm} + F_{u_x} u_{xm}) u^* dxdt$$

$$\int_t \int_\Omega F_{u_t} u_{tm} u^* dxdt = \int_\Omega [F_{u_t} u_m u^*]_0^T dx - \int_t \int_\Omega (F_{u_t} u^*)_t u_m dxdt$$

$$\int_t \int_\Omega F_{u_x} u_{xm} u^* dxdt = \int_t \int_\Gamma F_{u_x} u_m u^* dldt - \int_t \int_\Omega (F_{u_x} u^*)_x u_m dxdt$$

Example - Bringing All Together

Hence:

$$\int_t \int_\Omega \frac{dF}{dm} u^* dx dt$$

$$= \int_t \int_\Omega \{F_m u^* - u_m (F_u u^* - (F_{u_t} u^*)_t - (F_{u_x} u^*)_x)\} dx dt$$

Example - Bringing All Together

Hence:

$$\int_t \int_\Omega \frac{dF}{dm} u^* dx dt$$

$$= \int_t \int_\Omega \{F_m u^* - u_m (F_u u^* - (F_{u_t} u^*)_t - (F_{u_x} u^*)_x)\} dx dt$$

$$+ \int_\Omega [F_{u_t} u_m u^*]_0^T dx$$

Example - Bringing All Together

Hence:

$$\int_t \int_\Omega \frac{dF}{dm} u^* dx dt$$

$$= \int_t \int_\Omega \{F_m u^* - u_m (F_u u^* - (F_{u_t} u^*)_t - (F_{u_x} u^*)_x)\} dx dt$$

$$+ \int_\Omega [F_{u_t} u_m u^*]_0^T dx$$

$$+ \int_t \int_\Gamma F_{u_x} u_m u^* dl dt$$

Example - Defining the Adjoint Equation

Remember the gradient of the augmented cost function:

$$\frac{dL(x,u,m)}{dm} = \frac{dJ(x,u,m)}{dm} + \int_t \int_\Omega \frac{dF(u,u_t,u_x,m)}{dm} u^* dx dt$$

Example - Defining the Adjoint Equation

Remember the gradient of the augmented cost function:

$$\frac{dL(x,u,m)}{dm} = \frac{dJ(x,u,m)}{dm} + \int_t \int_\Omega \frac{dF(u,u_t,u_x,m)}{dm} u^* dx dt$$

This now becomes:

$$\frac{dL}{dm} = \int_t \int_\Omega \{F_m u^* - u_m (F_u u^* - (F_{u_t} u^*)_t - (F_{u_x} u^*)_x - j_u) + j_m\} dx dt$$

Example - Defining the Adjoint Equation

Remember the gradient of the augmented cost function:

$$\frac{dL(x,u,m)}{dm} = \frac{dJ(x,u,m)}{dm} + \int_t \int_\Omega \frac{dF(u,u_t,u_x,m)}{dm} u^* dx dt$$

This now becomes:

$$\frac{dL}{dm} = \int_t \int_\Omega \{F_m u^* - u_m (F_u u^* - (F_{u_t} u^*)_t - (F_{u_x} u^*)_x - j_u) + j_m\} dx dt$$

$$+ \int_\Omega [F_{u_t} u_m u^*]_0^T dx$$

$$+ \int_t \int_\Gamma F_{u_x} u_m u^* dl dt$$

Example - Boundary and Initial Cond's

Recall the last two terms:

$$\int_{\Omega} [F_{u_t} u_m u^*]_0^T dx$$

$$\int_t \int_{\Gamma} F_{u_x} u_m u^* dl dt$$

Example - Boundary and Initial Cond's

Recall the last two terms:

$$\int_{\Omega} [F_{u_t} u_m u^*]_0^T dx$$

$$\int_t \int_{\Gamma} F_{u_x} u_m u^* dl dt$$

The BC's should be set such that $u_x = 0$ on the boundary.

In general these two terms are treated through setting of the appropriate boundary and initial conditions and will be set aside for the rest of the derivation.

Example - Adjoint Equations

This leaves us with:

$$\frac{dL}{dm} = \int_t \int_\Omega \{F_m u^* - u_m(F_u u^* - (F_{u_t} u^*)_t - (F_{u_x} u^*)_x - j_u) + j_m\} dx dt$$

Example - Adjoint Equations

This leaves us with:

$$\frac{dL}{dm} = \int_t \int_\Omega \{F_m u^* - u_m (F_u u^* - (F_{u_t} u^*)_t - (F_{u_x} u^*)_x - j_u) + j_m\} dx dt$$

We now define the **adjoint equation** as follows:

$$F_u u^* - (F_{u_t} u^*)_t - (F_{u_x} u^*)_x = j_u$$

Example - Adjoint Equations

This leaves us with:

$$\frac{dL}{dm} = \int_t \int_\Omega \{F_m u^* - u_m (F_u u^* - (F_{u_t} u^*)_t - (F_{u_x} u^*)_x - j_u) + j_m\} dx dt$$

We now define the **adjoint equation** as follows:

$$F_u u^* - (F_{u_t} u^*)_t - (F_{u_x} u^*)_x = j_u$$

Hence the sensitivity of our model simplifies to:

$$\frac{dL}{dm} = \int_t \int_\Omega (F_m u^* + j_m) dx dt$$

Example - Adjoint Equations

We chose the following form of F for our example:

$$F = \frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x}$$

Example - Adjoint Equations

We chose the following form of F for our example:

$$F = \frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x}$$

Substituting for F in the adjoint equation:

$$F_u u^* - (F_{u_t} u^*)_t - (F_{u_x} u^*)_x = j_u$$

gives:

$$\frac{\partial u}{\partial x}u^* - \frac{\partial u^*}{\partial t} - u\frac{\partial u^*}{\partial x} - \frac{\partial u}{\partial x}u^* = \frac{\partial j}{\partial u}$$

$$-\frac{\partial u^*}{\partial t} - u\frac{\partial u^*}{\partial x} = \frac{\partial j}{\partial u}$$

Forward Sensitivity

The forward sensitivity equation is derived by differentiating F with respect to m:

$$\frac{dF}{dm} = \frac{\partial F}{\partial m} + \frac{\partial F}{\partial u}\frac{\partial u}{\partial m} + \frac{\partial F}{\partial u_t}\frac{\partial u_t}{\partial x} + \frac{\partial F}{\partial u_x}\frac{\partial u_x}{\partial m}$$

Forward and Adjoint

Forward sensitivity:

- shows the development of the initial perturbations of input variables
- needs to be evaluated for every parameter perturbation we choose to analyse

Forward and Adjoint

Forward sensitivity:

- shows the development of the initial perturbations of input variables
- needs to be evaluated for every parameter perturbation we choose to analyse

Adjoint sensitivity:

- finds the source of a specific anomaly
- does NOT model physical quantities
- models the *sensitivity* of a property to these quantities

Tangent linear vs Adjoint

Adjoint:

- efficiently computes sensitivity of a model with a large number of sensitivity parameters and few objective/cost functions
- simulates the development of sensitivities backwards in time
- can find origins of any anomaly

Tangent linear:

- efficiently computes sensitivity of a model with relatively few parameters or a large number of objective/cost functions
- simulates the development of perturbations with time

Limitations

- underlying functions non-differentiable
  - in ocean models: sub grid processes often formulated in a non-differentiable way

Limitations

- underlying functions non-differentiable
  - in ocean models: sub grid processes often formulated in a non-differentiable way
- based on linear approximation
  - sensitivities for nonlinear models only valid at a certain point in phase space
  - sensitivities for highly nonlinear (or chaotic) models might change rapidly with varying point of linearization —> useless results!

Adjoint methods - store or recalculate?

Relies on ability to efficiently store or recompute the simulation in the forward direction

- store all the data from the forward model
- re-run the entire model whenever the forward solution is needed
- store a subset of data and develop interpolation schemes
- checkpointing

Continuous or Discrete?

Continuous:

Nonlinear PDE - Linear PDE - Linear adjoint PDE - discrete adjoint

Discrete:

Nonlinear PDE - Nonlinear discrete PDE - Linear discrete equations - discrete adjoint equation

Continuous or Discrete?

Advantages of continuous derivation:

- highlights the physical significance of adjoint variables in boundary condition
- cheaper to derive than discrete method
- adaptive mesh models are naturally consistent with the adjoint equations - easier to implement

Continuous or Discrete?

Advantages of continuous derivation:

- highlights the physical significance of adjoint variables in boundary condition
- cheaper to derive than discrete method
- adaptive mesh models are naturally consistent with the adjoint equations - easier to implement

Advantages of discrete derivation:

- obtains exact gradient to discrete objective function, optimisation/sensitivity converges fully
- conceptually straight forward to create program

Continuous or Discrete?

Disadvantages of continuous derivation:

- possible problems with implementation of boundary conditions - can cause the adjoint models to become inadmissible for the chosen cost function
- the optimization/sensitivity might not converge

Continuous or Discrete?

Disadvantages of continuous derivation:

- possible problems with implementation of boundary conditions - can cause the adjoint models to become inadmissible for the chosen cost function
- the optimization/sensitivity might not converge

Disadvantages of discrete derivation:

- code is long
- inefficient

THE END